

#5

OICE

RAW SEQUENCE LISTING

DATE: 07/25/2001

PATENT APPLICATION: US/09/849,243

TIME: 16:19:01

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\07252001\I849243.raw

SEQUENCE LISTING

ENTERED

4 (1) GENERAL INFORMATION:
6 (i) APPLICANT: Kirschbaum, Bernd
7 Berglund, Erick
8 Meisterernst, Michael
9 Polites, Greg
11 (ii) TITLE OF INVENTION: PURIFICATION OF HIGHER ORDER TRANSCRIPTION
12 COMPLEXES FROM TRANSGENIC
13 NON-HUMAN ANIMALS
15 (iii) NUMBER OF SEQUENCES: 17
17 (iv) CORRESPONDENCE ADDRESS:
18 (A) ADDRESSEE: HELLER, EHRMAN, WHITE & McAULIFFE
19 (B) STREET: 1666 K Street, N.W., Suite 300
20 (C) CITY: Washington
21 (D) STATE: D.C.
22 (E) COUNTRY: USA
23 (F) ZIP: 20006
25 (v) COMPUTER READABLE FORM:
26 (A) MEDIUM TYPE: Floppy disk
27 (B) COMPUTER: IBM PC compatible
28 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
29 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
31 (vi) CURRENT APPLICATION DATA:
C--> 32 (A) APPLICATION NUMBER: US/09/849,243 BY
C--> 33 (B) FILING DATE: 07-May-2001 BY
35 (viii) ATTORNEY/AGENT INFORMATION:
36 (A) NAME: Granados, Patricia D.
37 (B) REGISTRATION NUMBER: 33,683
38 (C) REFERENCE/DOCKET NUMBER: 38005-0148
40 (ix) TELECOMMUNICATION INFORMATION:
41 (A) TELEPHONE: (202)912-2000
42 (B) TELEFAX: (202)912-2020
47 (2) INFORMATION FOR SEQ ID NO: 1:
49 (i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 12 amino acids
51 (B) TYPE: amino acid
52 (D) TOPOLOGY: linear
54 (ii) MOLECULE TYPE: peptide
58 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
60 Met Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Val
61 1 5 10
64 (2) INFORMATION FOR SEQ ID NO: 2:
66 (i) SEQUENCE CHARACTERISTICS:
67 (A) LENGTH: 11 amino acids
68 (B) TYPE: amino acid
69 (D) TOPOLOGY: linear
71 (ii) MOLECULE TYPE: peptide

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/849,243

DATE: 07/25/2001

TIME: 16:19:01

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\07252001\I849243.raw

75 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 77 Met Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
 78 1 5 10

81 (2) INFORMATION FOR SEQ ID NO: 3:

83 (i) SEQUENCE CHARACTERISTICS:

84 (A) LENGTH: 10 amino acids

85 (B) TYPE: amino acid

86 (D) TOPOLOGY: linear

88 (ii) MOLECULE TYPE: peptide

92 ~~(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:~~

94 Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala

95 1 5 10

98 (2) INFORMATION FOR SEQ ID NO: 4:

100 (i) SEQUENCE CHARACTERISTICS:

101 (A) LENGTH: 9 amino acids

102 (B) TYPE: amino acid

103 (D) TOPOLOGY: linear

105 (ii) MOLECULE TYPE: peptide

109 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

111 Tyr Pro Tyr Asp Val Pro Asp Tyr Ala

112 1 5

115 (2) INFORMATION FOR SEQ ID NO: 5:

117 (i) SEQUENCE CHARACTERISTICS:

118 (A) LENGTH: 22 base pairs

119 (B) TYPE: nucleic acid

120 (C) STRANDEDNESS: single

121 (D) TOPOLOGY: linear

123 (ii) MOLECULE TYPE: cDNA

126 (ix) FEATURE:

127 (A) NAME/KEY: exon

128 (B) LOCATION: 1..22

132 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

134 GGAGCAACCG CCTGCTGGGT GC

22

137 (2) INFORMATION FOR SEQ ID NO: 6:

139 (i) SEQUENCE CHARACTERISTICS:

140 (A) LENGTH: 21 base pairs

141 (B) TYPE: nucleic acid

142 (C) STRANDEDNESS: single

143 (D) TOPOLOGY: linear

145 (ii) MOLECULE TYPE: cDNA

148 (ix) FEATURE:

149 (A) NAME/KEY: exon

150 (B) LOCATION: 1..21

153 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

155 CCTGTGTTGC CTGCTGGGAC G

21

158 (2) INFORMATION FOR SEQ ID NO: 7:

160 (i) SEQUENCE CHARACTERISTICS:

161 (A) LENGTH: 21 base pairs

162 (B) TYPE: nucleic acid

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/849,243

DATE: 07/25/2001

TIME: 16:19:01

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\07252001\I849243.raw

```

163         (C) STRANDEDNESS: single
164         (D) TOPOLOGY: linear
166     (ii) MOLECULE TYPE: cDNA
169     (ix) FEATURE:
170         (A) NAME/KEY: exon
171         (B) LOCATION: 1..21
174     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
176 GGAGACTGAA GTTAGGCCAG C                               21
179 (2) INFORMATION FOR SEQ ID NO: 8:
181     (i) SEQUENCE CHARACTERISTICS:
182         (A) LENGTH: 76 base pairs
183         (B) TYPE: nucleic acid
184         (C) STRANDEDNESS: single
185         (D) TOPOLOGY: linear
187     (ii) MOLECULE TYPE: cDNA
190     (ix) FEATURE:
191         (A) NAME/KEY: exon
192         (B) LOCATION: 1..76
196     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
198 GCGGCACCAG GCCGCTGCTG TGATGATGAT GATGATGGCT GCTGCCCATG ACTGCGTAAT   60
200 GCGGTCATGA CGCTTT                                           76
203 (2) INFORMATION FOR SEQ ID NO: 9:
205     (i) SEQUENCE CHARACTERISTICS:
206         (A) LENGTH: 75 base pairs
207         (B) TYPE: nucleic acid
208         (C) STRANDEDNESS: single
209         (D) TOPOLOGY: linear
211     (ii) MOLECULE TYPE: cDNA
214     (ix) FEATURE:
215         (A) NAME/KEY: exon
216         (B) LOCATION: 1..75
219     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
221 GAAGGGGGTG GGGGAGGCAA GGGTACATGA GAGCCATTAC GTCGTCTTCC TGAATCCCTT   60
223 TAGCCGCTTT GCTCG                                           75
226 (2) INFORMATION FOR SEQ ID NO: 10:
228     (i) SEQUENCE CHARACTERISTICS:
229         (A) LENGTH: 22 base pairs
230         (B) TYPE: nucleic acid
231         (C) STRANDEDNESS: single
232         (D) TOPOLOGY: linear
234     (ii) MOLECULE TYPE: cDNA
237     (ix) FEATURE:
238         (A) NAME/KEY: exon
239         (B) LOCATION: 1..22
242     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
244 CCCTATGACG TCCCGGATTA CG                                   22
247 (2) INFORMATION FOR SEQ ID NO: 11:
249     (i) SEQUENCE CHARACTERISTICS:
250         (A) LENGTH: 22 base pairs

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/849,243

DATE: 07/25/2001

TIME: 16:19:01

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\07252001\I849243.raw

```

251      (B) TYPE: nucleic acid
252      (C) STRANDEDNESS: single
253      (D) TOPOLOGY: linear
255      (ii) MOLECULE TYPE: cDNA
258      (ix) FEATURE:
259          (A) NAME/KEY: exon
260          (B) LOCATION: 1..22
263      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
265 GTGGAGTGGT GCCCGGCAAG GG
268 (2) INFORMATION FOR SEQ ID NO: 12:
270      (i) SEQUENCE CHARACTERISTICS:
271          (A) LENGTH: 19 amino acids
272          (B) TYPE: amino acid
273          (D) TOPOLOGY: linear
275      (ii) MOLECULE TYPE: peptide
279      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
281      Met Gly Ser Ser His His His His Ser Ser Gly Leu Val Pro
282      1          5          10          15
284      Arg Gly Cys
288 (2) INFORMATION FOR SEQ ID NO: 13:
290      (i) SEQUENCE CHARACTERISTICS:
291          (A) LENGTH: 1310 base pairs
292          (B) TYPE: nucleic acid
293          (C) STRANDEDNESS: single
294          (D) TOPOLOGY: linear
296      (ii) MOLECULE TYPE: cDNA
299      (ix) FEATURE:
300          (A) NAME/KEY: exon
301          (B) LOCATION: 1..1310
304      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
306 CCATGGGCTA TCCCTATGAC GTCCCGGATT ACGCAGTCAT GGGCAGCAGC CATCATCATC 60
308 ATCATCACAG CAGCGGCCTG GTGCCGCGCG GCAGCCATAT GGATCAGAAC AACAGCCTGC 120
310 CACCTTACGC TCAGGGCTTG GCCTCCCCTC AGGGTGCCAT GACTCCCGGA ATCCCTATCT 180
312 TTAGTCCAAT GATGCCTTAT GGCAGTGGAC TGACCCACAC GCCTATTTCAG AACACCAATA 240
314 GTCTGTCTAT TTTGGAAGAG CAACAAAGGC AGCAGCAGCA ACAACAACAG CAGCAGCAGC 300
316 AGCAGCAGCA GCAGCAACAG CAACAGCAGC AGCAGCAGCA GCAGCAGCAG CAGCAGCAGC 360
318 AGCAGCAGCA GCAGCAGCAA CAGGCAGTGG CAGCTGCAGC CGTTCAGCAG TCAACGTCCC 420
320 AGCAGGCAAC ACAGGGAACC TCAGGCCAGG CACCACAGCT CTTCCACTCA CAGACTCTCA 480
322 CAACTGCACC CTTGCCGGGC ACCACTCCAC TGTATCCCTC CCCCATGACT CCCATGACCC 540
324 CCATCACTCC TGCCACGCCA GCTTCGGAGA GTTCTGGGAT TGTACCGCAG CTGCAAAATA 600
326 TTGTATCCAC AGTGAATCTT GGTGTGTAAC TTGACCTAAA GACCATTGCA CTTCGTGCCC 660
328 GAAACGCCGA ATATAATCCC AAGCGGTTTG CTGCGGTAAT CATGAGGATA AGAGAGCCAC 720
330 GAACCACGGC ACTGATTTTC AGTTCCTGGG AAATGGTGTG CACAGGAGCC AAGAGTGAAG 780
332 AACAGTCCAG ACTGGCAGCA AGAAAATATG CTAGAGTTGT ACAGAAGTTG GGTTTTCAG 840
334 CTAAGTTCTT GGACTTCAAG ATTCAAGACA TGGTGGGGAG CTGTGATGTG AAGTTTCCTA 900
336 TAAGGTTAGA AGGCCTTGTG CTCACCCACC AACAATTTAG TAGTTATGAG CCAGAGTTAT 960
338 TTCCTGGTTT AATCTACAGA ATGATCAAAC CCAGAATTGT TCTCCTTATT TTTGTTTCTG 1020
340 GAAAGTTGT ATTAACAGGT GCTAAAGTCA GAGCAGAAAT TTATGAAGCA TTTGAAAACA 1080
342 TCTACCCTAT TCTAAAGGGA TTCAGGAAGA CGACGTAATG GCTCTCATGT ACCCTTGCCT 1140

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/849,243

DATE: 07/25/2001

TIME: 16:19:01

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\07252001\I849243.raw

```

344 CCCCCACCCC CTTCTTTTTT TTTTTTAAA CAAATCAGTT TGTTTTGGTA CCTTTAAATG      1200
346 GTGGTGTGTG GAGAAGATGG ATGTTGAGTT GCAGGGTGTG GCACCAGGTG ATGCCCTTCT      1260
348 GTAAGTGCCC CTTCCGGCAT CCCGGAATTC CTGCAGCCCA ACGCGGCCGC      1310
351 (2) INFORMATION FOR SEQ ID NO: 14:
353     (i) SEQUENCE CHARACTERISTICS:
354         (A) LENGTH: 4286 base pairs
355         (B) TYPE: nucleic acid
356         (C) STRANDEDNESS: single
357         (D) TOPOLOGY: linear
359     (ii) MOLECULE TYPE: cDNA
362     (ix) FEATURE:
363         (A) NAME/KEY: exon
364         (B) LOCATION: 1..4286
367     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
369 GAATTCCTCT GCAGGTCAGT TAGCGTTGGC CACATAGTAG GTTCTCAAAT ACTTGTTAAT      60
371 AAATAAGTTT GTTCGAGAAG CTGGGCAATG ATATTCTACA GCTGGAAGAA GAAACATAAT      120
373 GATCTAGTAA TTAGCTCAAT TAAAAATAAA CGTTCTTCTT TCCTCAGAGG AGCATTTCCT      180
375 AAGGCCTGCC TTGATAGCCA TCCAAAAAGG CCAAGCTCAT CCAATCTTGC CCTAGATTTA      240
377 TGCTAAAATG CAGTTACAAT CGATAGGATG ACAGAAAACG ACAGCACTTA TTAAATATA      300
379 ATAGGCACTT ATTTAAATAG GAGAAGCTGT GACTTCATAG CAAGTGTTGG GGTTAGGAAA      360
381 CTGGGTGGAT AAAGTGTCTG ATGCTGTAGA TCTTAGCCTC TACATGAGAT CATGTGGAAA      420
383 ATCTGAAAGC ATTTAGGTTT CCTTATGTTT GCAATCAAAT AACTGTACAC CTTTAAATTT      480
385 AAAAAGTACC ATGAGGCACA CACAACACT CGCAGGAAGT TTTTGGCGTA ACAAACACTAG      540
387 AATTAGATCT AAAAGCTAAC TGTAGGACTG AGTCTATTCT AAAGTAAAG CCTGGACATC      600
389 TGGAGTACCA GGGGGAGATG ACGTGTTACG GGCTTCCATA AAAGCAGCTG GCTTTGAATG      660
391 GAAGGAGCCA AGAGGCCAGC ACAGGAGCGG ATTCGTCGCT TTCACGGCCA TCGAGCCGAA      720
393 CCTCTCGCAA GTCCGTGAGC CGTTAAGGAG GCGCCAGTCC CCGACCCTTC GCGCCAGCC      780
395 CCTCGGGGTC CCCGGGCGTG GTACTCCTTG CCACACGGGA GGGGCGCGGA AGCCGGGGCG      840
397 GAGGAGGAGC CAACCCCGGG CTGGGCTGAG ACCCGCAGAG GAAGACGCTC TAGGGATTG      900
399 TCCCGGACTA GCGAGATGGC AAGGCTGAGG ACGGGAGGCT GATTGAGAGG CGAAGGTACA      960
401 CCCTAATCTC AATACAACCT TTGGAGCTAA GCCAGCAATG GTAGAGGGAA GATTCTGCAC      1020
403 GTCCCTTCCA GCGCGCCTCC CCGTCACCAC CCCCCCAAC CCGCCCCGAC CGGAGCTGAG      1080
405 AGTAATTCAT ACAAAGGAC TCGCCCCTGC CTTGGGGAAT CCCAGGGACC GTCGTTAAAC      1140
407 TCCCACTAAC GTAGAACCCA GAGATCGCTG CGTTCCCGCC CCCTCACCCG CCCGCTCTCG      1200
409 TCATCACTGA GGTGGAGAAG AGCATGCGTG AGGCTCCGGT GCGCGTCAGT GGGCAGAGCG      1260
411 CACATCGCCC ACAGTCCCCG AGAAGTTGGG GGGAGGGGTC GGCAATTGAA CCGGTGCCTA      1320
413 GAGAAGGTGG CGCGGGGTAA ACTGGGAAAG TGATGTCGTG TACTGGCTCC GCCTTTTCC      1380
415 CGAGGGTGGG GGAGAACCGT ATATAAGTGC AGTAGTCGCC GTGAACGTTT TTTTTCGCAA      1440
417 CGGGTTTGCC GCCAGAACAC AGGTAAGTGC CGTGTGTGGT TCCCGCGGGC CTGGCCTCTT      1500
419 TACGGGTTAT GGCCCTTGCG TGCCTTGAAT TACTTCCACG CCCCTGGCTG CAGTACGTGA      1560
421 TTCTTGATCC CGAGCTTCGG GTTGGAAGTG GGTGGGAGAG TTCGAGGCCT TGCGCTTAAG      1620
423 GAGCCCTTTC GCCTCGTGCT TGAGTTGAGG CCTGGCCTGG GCGCTGGGGC CGCCGCGTGC      1680
425 GAATCTGGTG GCACCTTCGC GCCTGTCTCG CTGCTTTTGA TAAGTCTCTA GCCATTTAAA      1740
427 ATTTTGTATG ACCTGCTGCG ACGCTTTTTT TCTGGCAAGA TAGTCTTGTA AATGCGGGCC      1800
429 AAGATCTGCA CACTGGTATT TCGGTTTTTG GGGCCGCGGG CGGCGACGGG GCCCGTGCGT      1860
431 CCCAGCGCAC ATGTTGCGCG AGGCGGGGCC TGCGAGCGCG GCCACCGAGA ATCGGACGGG      1920
433 GGTAGTCTCA AGCTGGCCGG CCTGCTCTGG TGCCTGGCCT CGCGCCGCCG TGTATCGCCC      1980
435 CGCCCTGGGC GGCAAGGCTG GCGCGGTCGG CACCAAGTTG GTGAGCGGAA AGATGGCCGC      2040
437 TTCCCGGCCG TGCTGCAGGG AGCTCAAAAT GGAGGACGCG GCGCTCGGGA GAGCGGGCGG      2100

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/849,243

DATE: 07/25/2001

TIME: 16:19:02

Input Set : A:\PTO_VSK.txt

Output Set: N:\CRF3\07252001\I849243.raw

L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]